

# 12th Annual Acquisition Research Symposium

## Big Data: Big Confusion? Big Challenges?

---

Mary Maureen Brown, Ph.D.  
UNC Charlotte  
([marbrown@uncc.edu](mailto:marbrown@uncc.edu))

| Report Documentation Page  |                                    |                                     |   | Form Approved<br>OMB No. 0704-0188                  |                                 |
|--|------------------------------------|-------------------------------------|---|---|---------------------------------|
| Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. |                                    |                                     |   |   |                                 |
| 1. REPORT DATE<br><b>MAY 2015</b>  |                                    | 2. REPORT TYPE                      |   | 3. DATES COVERED<br><b>00-00-2015 to 00-00-2015</b> |                                 |
| 4. TITLE AND SUBTITLE<br><b>Big Data: Big Confusion? Big Challenges?</b>   |                                    |                                     |   | 5a. CONTRACT NUMBER                                 |                                 |
|  |                                    |                                     |   | 5b. GRANT NUMBER                                    |                                 |
|  |                                    |                                     |   | 5c. PROGRAM ELEMENT NUMBER                          |                                 |
| 6. AUTHOR(S)   |                                    |                                     |   | 5d. PROJECT NUMBER                                  |                                 |
|  |                                    |                                     |   | 5e. TASK NUMBER                                     |                                 |
|  |                                    |                                     |   | 5f. WORK UNIT NUMBER                                |                                 |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br><b>University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC, 28223</b>  |                                    |                                     |   | 8. PERFORMING ORGANIZATION REPORT NUMBER            |                                 |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)  |                                    |                                     |   | 10. SPONSOR/MONITOR'S ACRONYM(S)                    |                                 |
|  |                                    |                                     |   | 11. SPONSOR/MONITOR'S REPORT NUMBER(S)              |                                 |
| 12. DISTRIBUTION/AVAILABILITY STATEMENT<br><b>Approved for public release; distribution unlimited</b>  |                                    |                                     |   |   |                                 |
| 13. SUPPLEMENTARY NOTES<br><b>Presented at the 12th Annual Acquisition Research Symposium held May 13-14, 2015 in Monterey, CA.</b>  |                                    |                                     |   |   |                                 |
| 14. ABSTRACT   |                                    |                                     |   |   |                                 |
| 15. SUBJECT TERMS  |                                    |                                     |   |   |                                 |
| 16. SECURITY CLASSIFICATION OF:  |                                    |                                     | 17. LIMITATION OF ABSTRACT<br><b>Same as Report (SAR)</b> | 18. NUMBER OF PAGES<br><b>14</b>                    | 19a. NAME OF RESPONSIBLE PERSON |
| a. REPORT<br><b>unclassified</b>   | b. ABSTRACT<br><b>unclassified</b> | c. THIS PAGE<br><b>unclassified</b> |   |   |                                 |

National  
Consortium for  
Data Science



the UNIVERSITY of NORTH CAROLINA  
General Administration



THE UNIVERSITY  
of NORTH CAROLINA  
at CHAPEL HILL



NC STATE UNIVERSITY



THE UNIVERSITY of NORTH CAROLINA  
GREENSBORO



Data Science Initiative

Belk College of Business / College of Computing and Informatics /  
College of Health and Human Services

Search this site

Home

Contact Us

Academic Advising



MCNC  
Connecting North Carolina's Future Today



EMC<sup>2</sup>



Deloitte.

Bank of America



belk

sas

THE  
POWER  
TO KNOW.



Carolina's HealthCare System

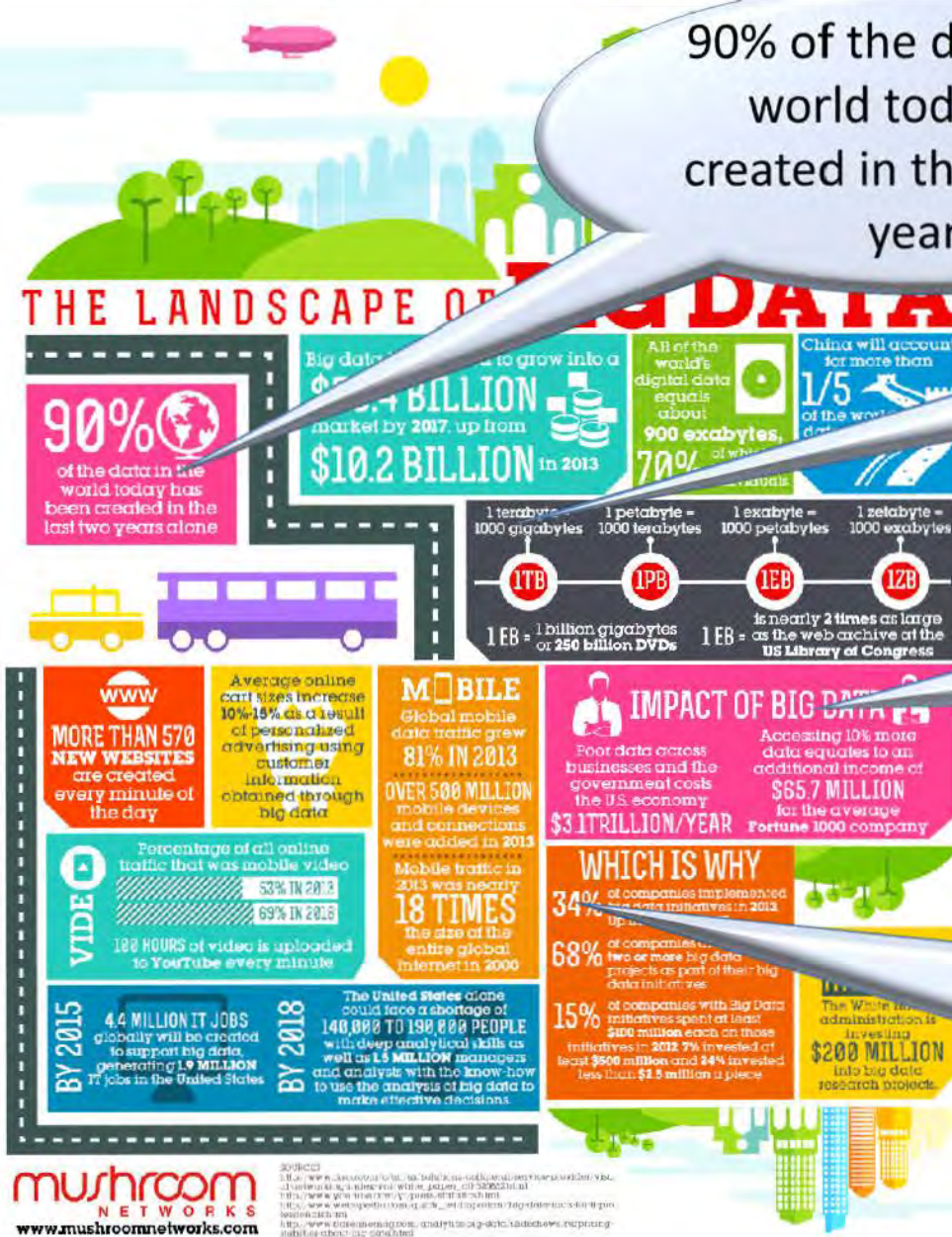


90% of the data in the world today was created in the last two years

Big Data growth from \$10 Billion in 2013 to \$53 Billion by 2017

Poor data across the private and public sectors costs the U.S. economy \$3 Trillion/year

By 2018, shortage of 190,000 people with deep analytical skills & 1.5 million managers and analysts evidence based decisions



# Value of Big Data

Improvements in strategic decision making operational efficiency

- Generating new data sources
  - Predictive analytics
  - Market research

(AMA Survey top 5)

# Data to Decisions: “placing a big bet on big data”

The Department of Defense investing \$250 million annually

- ☐ Harness and utilize massive data in new ways and bring together sensing, perception and decision support to make truly autonomous systems that can maneuver and make decisions on their own.
- ☐ Improve situational awareness to help warfighters and analysts and provide increased support to operations.

*Today, no matter what business you are in, technology, problem-solving and data analytics are at the heart of it.*

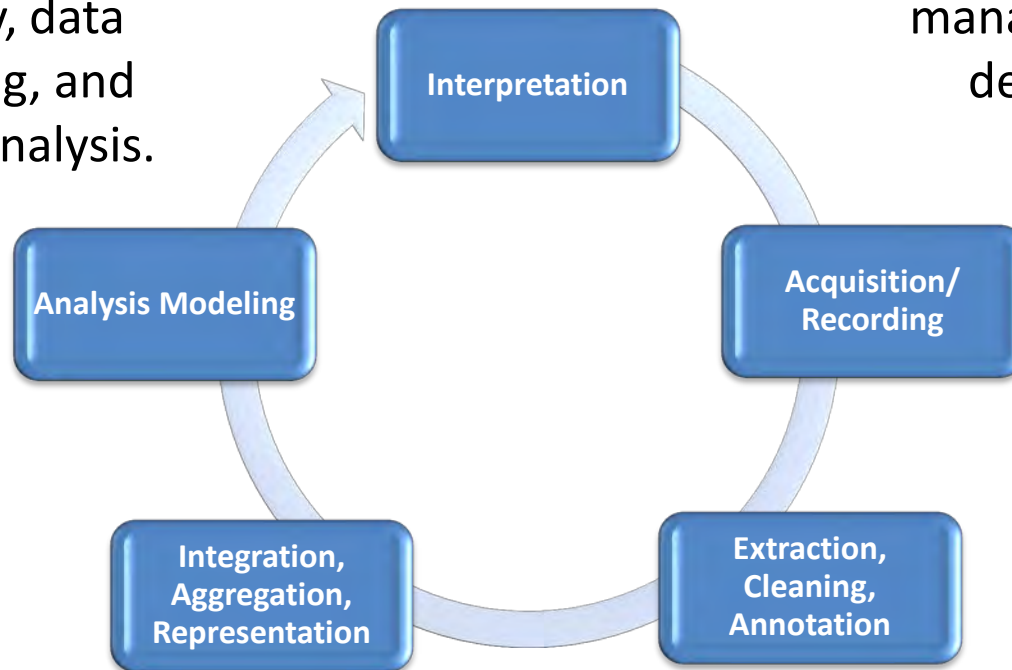
# Definitions

## Data Science & Business Analytics

Data Science - knowledge extraction, information discovery, information harvesting, data archaeology, data pattern processing, and exploratory data analysis.



Business Analytics – heavily grounded in OR, including explanatory and predictive modeling, and fact based management to drive decision making



## Big Datasets Bring....

noise accumulation,  
spurious correlations and  
incidental homogeneity

heavy computational cost  
and algorithmic instability

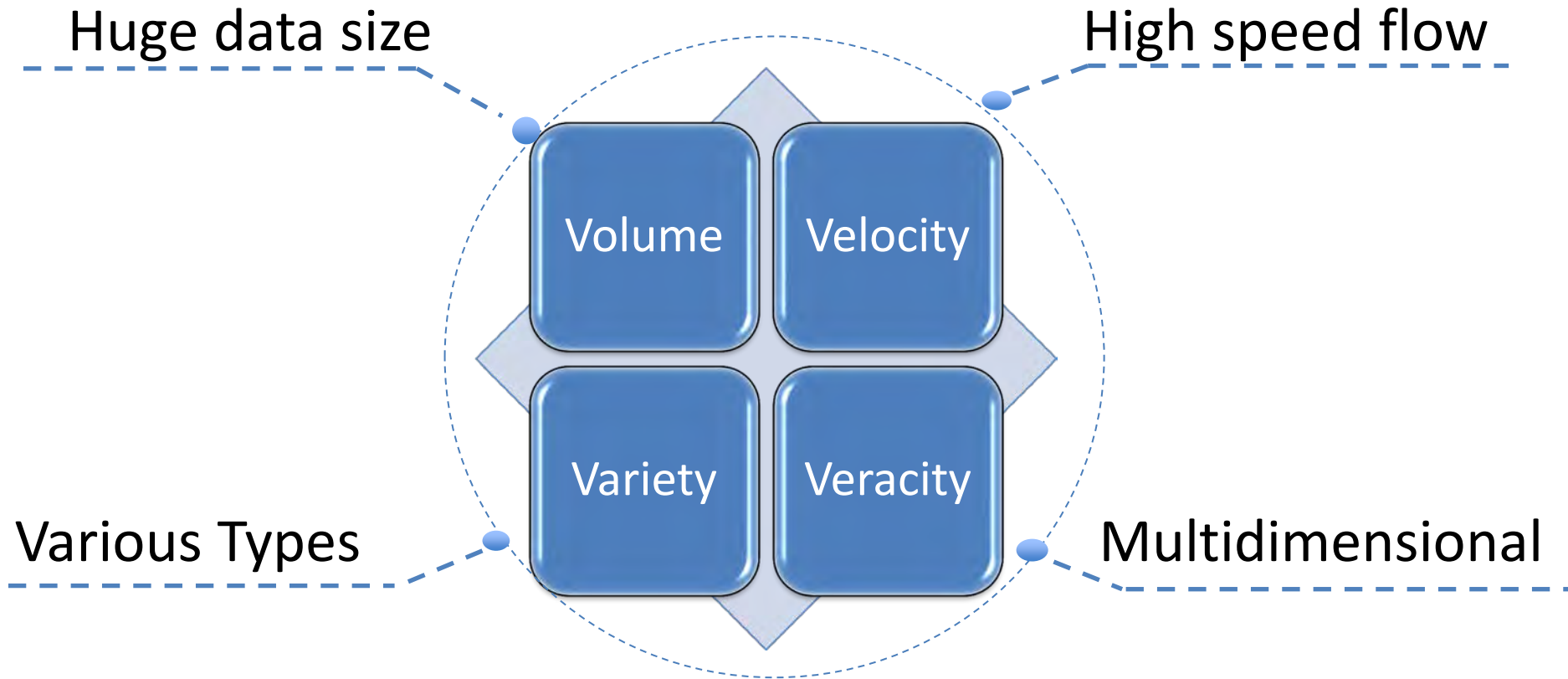
Aggregation from different  
systems can result in  
statistical biases &  
measurement error

## Big Datasets Demand....

**more  
adaptive and  
robust  
procedures**



According to IBM,  
2.5 quintillion bytes of data are produced every day



*Big data will intensify the need:*

- For changes in data quality and governance,
- For embedding analytics into operational systems, and
- For security, privacy, and regulatory compliance

Big Data's power does not erase the need for human insight  
– it only elevates it.

Typical software  
engineering problems

Four management challenges:

1. Leadership
2. Talent management
3. Decision making responsibility and accountability
4. Culture

Far and above all others

*Turning data into insight is by far the biggest challenge...Forcing a major change in paradigms*

Traditional Analytical Approach – Specifications and Requirements for canned queries and reports

*Too many permutations*

Big Data Analytical Approach – Foster On Demand Analysis

Skills gap - Most of us today are woefully data illiterate

### Top 5 Needs:

- Critical Thinking
- Problem Solving
- Extrapolating Conclusions
- Communicating and Presenting
- Evidence Based Decision Making





## Big Data, 'Breaking Bad' And Orange Juice

MailChimp's chief data scientist has a unique way of explaining data science to the masses. If you know a little Excel, you're ready to learn.

What's the best way to teach data science to people who lack sufficient training in analytics, computer science, modeling and statistics? A little humor can't hurt, perhaps, particularly when delving into potentially mind-numbing topics like algorithms and programming.

[John Foreman](#), chief data scientist for MailChimp, an email marketing service provider, is doing his part to demystify the dark art of data science by speaking and writing extensively on the subject. His [data science blog](#) features an eclectic blend of real world -- if a bit unorthodox -- examples of big data in



**Big Data Analytics  
Masters Degrees: 20  
Top Programs**

(click image for larger view and for

“What's the best way to teach data science to people who lack sufficient training in analytics, computer science, modeling and statistics?”

“A little humor can't hurt”

# Take Away

“The evidence is clear: Data-driven decisions  
tend to be better decisions”

## Biggest Challenge

organizations are in the difficult position of  
having to build the capacity to approach  
problems in an analytical way

Erik Brynjolfsson

Andrew McAfee